Economic instruments and institutional constraints: possible schemes for SO2 emissions trading in the EU

Olivier Godard

Publication in S. Sorrell and J. Skea (eds.) *Pollution for Sale – Emissions trading and joint implementation*.

Cheltenham, Edward Elgar, 1999, pp. 272-295

15. Economic instruments and institutional constraints: possible schemes for SO_2 emissions trading in the EU

Olivier Godard[1 2]

1. INTRODUCTION

Following the 1979 Geneva Convention on Long Range Transboundary Air Pollution (LRTAP), two protocols on sulphur emissions have been adopted. The 1985 Helsinki Protocol established a uniform abatement target of 30 per cent for all Parties, on the basis of 1980 emissions, while the 1994 Oslo Protocol defined new national targets for SO_2 abatement, which for the first time are differentiated across countries. The Oslo Protocol also fixed emission standards for new sources, and agreed on the specifications of the best available technologies to be used by operators. The agreed long term objective is to reduce SO_2 emissions so as to respect critical loads[3] for acid deposition everywhere in Europe[4]. However, the regulatory requirements agreed upon within the Protocol are insufficient to achieve this

1. Senior researcher at CNRS.

2 This paper benefits from the results of a 1996 study conducted by the author and Christine Cros, research assistant at CIRED, for the DGII 'Economic and Financial Affairs' of the European Commission. See Cros and Godard (1998). Financial support from the French ADEME and Ministry of the Environment is also gratefully acknowledged.

3. Critical loads are defined as the maximum levels of acid deposition below which, according to current scientific knowledge, no significant damage to sensitive ecosystems can be demonstrated. Critical deposition for one zone is the maximum deposition compatible with the critical loads of specific ecosystems and land use within the zone. The 5-percentile critical loads correspond to deposition levels generating no significant damage for at least 95 per cent of the ecosystems within the area.

4. In the context of this paper, each time a territorial dimension is implied, 'Europe' or 'European' should be understood as the whole European territory covered by the Oslo Protocol of the Geneva Convention, i.e. including countries outside the EU. Quite evidently, then, EU refers to the territory of the present member countries of the EU.

long-term objective. To reduce the gap, it will be necessary to introduce new policy instruments or significantly tightening existing ones.

The purpose of this paper is to develop possible organisational schemes for SO_2 emissions trading within the EU[5] as a means of achieving a cost-effective trajectory to reach the long run objective. Under good conditions of information and organisation, a tradable permit system may enforce an overall environmental constraint while minimising the total abatement cost. This is why this instrument can at the same time be viewed as bringing an environmental and an economic improvement, when compared to administrative regulatory approaches focused on emission rates (Klaassen, 1996). To be viable, any proposed scheme must prove politically acceptable (distributive palatability), achieve a predictable improvement in environmental quality (environmental effectiveness), and most importantly be consistent with pre-existing basic rules and requirements (institutional acceptability) (Godard, 1995). The latter condition is the central issue addressed in this chapter.

The Oslo Protocol imposes one main constraint on SO_2 emissions: a set of individual national targets to be achieved at three dates in the future: 2000, 2005 and 2010. This choice has been guided by European-wide model optimisation exercises aiming at achieving an intermediate objective of a 60 per cent reduction in the gap between current levels of acid deposition and the 5-percentile critical loads for each 150km by 150km 'grid cell' within Europe. The Protocol also includes a provision that two parties could be authorised to join their abatement efforts. The detailed rules for this joint implementation concept have still to be determined by the executive body of the Convention. It seems that any such proposal would only be accepted if the parties can demonstrate that it will positively contribute to the long term environmental objective, i.e. improving environmental quality in the whole territory covered by the Convention. In the context of this paper, I have interpreted this requirement as an implicit second constraint, one submitting any SO_2 allowance trading scheme to the obligation of meeting a percentage abatement target regarding the gap between current levels of deposition and critical loads. The final 2010 target should be a 60 per cent reduction in this gap. However, to avoid disequilibrium between deposition and emission constraints, the deposition target may evolve with the same target years as the national emission ceilings. For example, we could assume abatement objectives of 55% in 2000, 57% in 2005 and 60% in 2010.

For the design of a trading scheme, this amounts to having two constraints to satisfy: overall national targets related to emissions; and the percentage abatement target as regards deposition in excess of critical loads. For areas where deposition is below critical loads, the second constraint is interpreted as ensuring critical loads are not exceeded. Beyond 2010, more stringent objectives (say 75 per cent or 90 per cent reduction in exceedance levels) could be adopted in a multi-phased approach.

Whatever the level chosen, there is no reason why these two types of constraints should automatically coincide. A key feature of this paper is to address this specific issue of satisfying two types of constraints when developing an allowance trading scheme. The joint

5. The acronym EU is also intended to cover the European Economic Community for years before the establishment of the EU.

implementation of the first constraint provides a global European cap on emissions, leading to allowances that can be directly expressed in quantities of emissions. The second constraint leads to allowances expressed in quantities of deposition by *unit zone* (see Box 1).

Box 1 Terminology for zones

The paper uses several definitions of geographical zones:

- *Deposition zone* is a general term for an area receiving acid deposition.
- *Unit zones* are the smallest zones for which critical loads have be determined. Under the UNECE protocols, these correspond to 150km by 150km squares.
- *Macro zones* are a grouping of several unit zones on the basis of some rule.
- *Trading zones* are areas within which trading is freely authorised with a one to one exchange rate. Between trading zones, trading may be forbidden or may be allowed with a specific exchange rate (for example 1:1.5). Trading zones may correspond to individual unit zones or a larger grouping of unit zones (a macro zone).

Two families of solutions are then considered:

- A system of two types of tradable allowances working in parallel. Firms have to gather the same amount of the two different types of allowances to obtain a permit to emit a corresponding amount of SO_2.
- An integrated system in which a single allowance type embodies both emission and deposition constraints. Two variants of an integrated system are considered which put different emphasis on the incentive mechanisms imposed on firms and national governments.

Such solutions may be judged rather complex to operate, more complex than those implemented by the Acid Rain Program in the US, for example. Nevertheless, the source of complication is to be found, not in allowance trading as such, but in the existing framework of the LRTAP and subsequent Protocols. To achieve any further progress towards the long-term goal it will be necessary to address the complication of the transition, whatever the policy instrument chosen. To offer the required level of guarantee, a ‘command and control’ approach would have to become either excessively stringent or cumbersome. I contend that by lowering the total cost of abatement, allowance trading may help facilitate an overall progression towards the abatement of acid deposition, even if progress cannot be guaranteed at the same pace for each unit zone of the European territorial grid. One major contribution of SO_2 allowance trading could be in the flexibility of timing given to the participating companies. With tradable allowances, firms are given the opportunity to optimise the timing of their investment decisions. This saves capital costs and encourages the efficient and timely adoption of technical innovations.

Approximately 65 per cent of total SO_2 emissions in the EU originate from the generation of electricity. It therefore seems quite natural to begin to implement trading schemes within this sector so as to gain practical experience of the tool within the EU. An opposite view would stress the low level of competition in the power generation sector, since oligopolies or monopolies are common structures in many EU countries. However, recent EU initiatives regarding the introduction of competition in this sector will increase the sensitivity to price signals and market opportunities. Positive synergies may then exist between electricity market liberalisation and allowance trading. Further extension of the trading scheme could be advantageously envisaged to oil refineries and to all industrial combustion plant above a certain size (e.g. 25MW).

The remainder of the paper is structured as follows. Section 2 provides an overview of some key variables for the design of a trading scheme, while section 3 discusses the issue of defining trading zones. Section 4 presents a system of two simultaneous, coupled allowance trading mechanisms. Section 5 is devoted to an integrated scheme including an auctioned market for 'unusable allowances', while section 6 describes an alternative integrated system incorporating compensations for member states. The main conclusions are summarised in section 7.

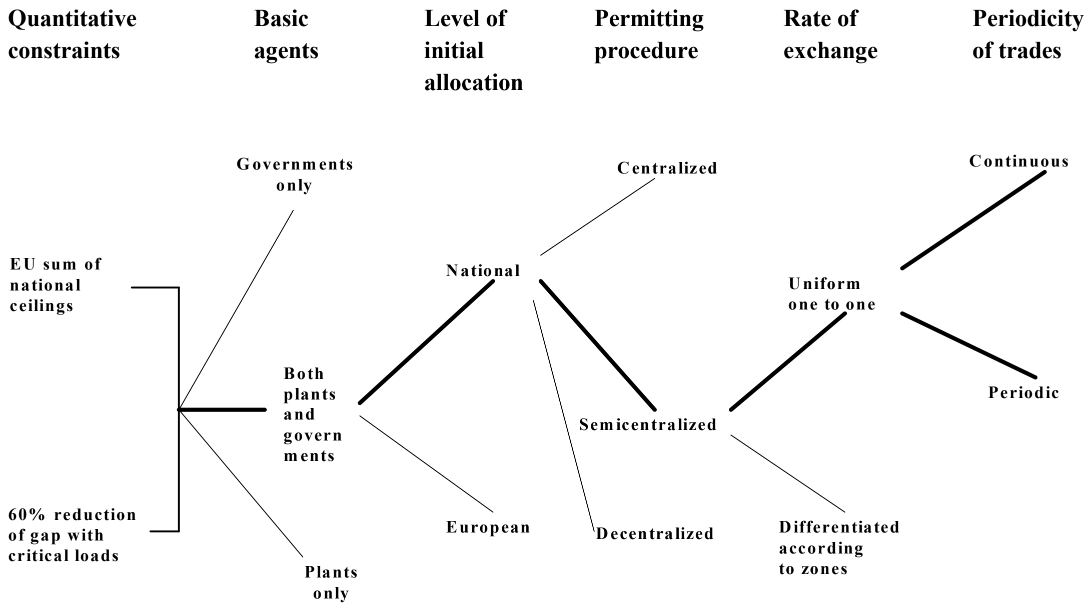
2. AN OVERVIEW OF SOME KEY DESIGN VARIABLES FOR TRADING SCHEMES

The success of a trading scheme depends crucially on the details of its design. Here, the following variables are selected for particular attention:

- *Basic agents*: which parties receive allowances and participate in trade.
- *Level of initial allocation:* whether this should be carried out at a European or member state level.
- *Permitting procedure*: how trades should be authorised to ensure that deposition constraints are not exceeded.
- *Exchange rate*: how emissions from widely separated sources may be compared to achieve approximate equivalence in environmental impact through trading.
- *Periodicity:* whether parties should be allowed to trade continuously or at periodic intervals.

Figure 1 summarises the choices made for these key variables, while subsequent sections explain these choices in more detail.

Figure 1 Key design variables for a EU scheme of SO_2 trading



2.1 Who trades?

Two types of actors could legitimately participate in a trading scheme: governments and electricity generators responsible for SO_2 emissions. An international trading mechanism that only operates between governments would leave significant opportunities for cost saving untapped. Because they directly control emissions and have the best access to technical and economic information, firms should be able to engage directly in a trading scheme. However a trading scheme organised solely at the firm level would correspond to a level of political integration which has yet to be achieved by the EU. It is EU member states which have taken on legally binding quantitative obligations under the Oslo Protocol and any trading arrangement must respect these obligations.

In this regard, the Oslo Protocol can be seen as a compromise between an agreement among independent States and a more integrated approach that could be developed if all States chose to behave as the members of a single political community. The EU context of negotiation and decision-making also looks intermediate: a network of regular co-operative links has been created and some acceptance of sectoral asymmetries in obligations and burden-sharing has been institutionalised.

The political balance of this compromise is reflected in the basic constraints embodied in the Oslo Protocol. Establishing emission ceilings on a country basis is a response to the first component of the compromise. So, even for a EU wide allowance trading scheme, the initial allocation of allowances at the plant level should be the responsibility of the individual member states. But the preliminary drafts of the Protocol were conceived with reference to the second component, with the idea of an integrated optimal plan for acid deposition for the whole European territory concerned, a plan sensitive to the location of deposition. This would require trading by individual sources.

These considerations lead to the choice of a *two-level* system, in which there is trading at both the government level and the plant level. Governments retain responsibility for the initial allocation of allowances to plants, together with the legal obligation to meet national emission ceilings. The allocation should provide a predictable framework in which individual plants can engage in trade to improve economic efficiency. Deposition constraints are enforced by the authorisation procedures for plant-to-plant trades, which are administered by the central authorities. These are described more fully below.

This choice is consistent with the concern for economic efficiency expressed by the EU as well as the Oslo Protocol. From the viewpoint of the economics of information, the potential for cost-effectiveness can only be exploited by giving appropriate incentives to decentralised management units; that is, to those actors who can most easily obtain the appropriate information concerning the available abatement opportunities, technologies and costs. This is what allowance trading is intended to achieve. If governments were to be considered as the only agents of the system, they would be unable to obtain some of the necessary information to minimise the social cost of abatement.

2.2 Which permitting procedure?

Permitting refers to the authorisation procedure for individual plant-to-plant trades. This must ensure that the deposition constraints are not exceeded. There are three broad options:

- A decentralized approach with free trade, i.e. without a specific authorisation procedure but according to agreed rules. Typically, trading may only be allowed within the same deposition zone, or between various zones on the basis of a fixed set of exchange rates.
- A semi-centralized approach based on physical modelling of the net impact of each proposed trade on acid deposition in each zone.
- A centralized or planned approach, in which physical and economic models are used to identify all possible beneficial trades compatible with the current target[6]. In order to be authorised, a projected trade should fit the pre-existing list of advantageous trades. Modelling is not used to authorise each proposal, but is used once, at the beginning of the period, to identify transactions that would be attractive and compatible.

Here, I suggest that a semi-centralised permitting system is the best option. The integrated assessment models used during negotiation of the Oslo Protocol combine physical modelling of the emission, transportation and deposition of pollutants with economic modelling of

6 This approach has been advocated by van Ierland et al. (1994).

abatement technologies and costs. However, since the cost functions used in the models are based on national aggregate values, the economic information is insufficient to identify the best opportunities for minimising abatement costs. This undermines the case for centralised permitting procedures. Instead, it is suggested that the economic part of the modelling exercise is put aside, leaving only the physical models of emission, transportation and deposition of pollutants together with the link between deposition and critical loads. The economic dimension of the allocation will be left to the decentralised calculations of individual sources. Sources will have to compute their own strategy and look for cost effective opportunities to trade.

In this case, the process is as follows:

- the Secretariat enters the distribution of all the allowances resulting from the national implementation of emission ceilings and technological standards into the physical assessment models;
- after a search period, two sources interested in trading find each other and agree on a trade proposal; they submit it to the Secretariat;
- the projected change in the location of emissions is entered in the agreed models which are run to provide forecasts about the environmental impacts for each deposition zone;
- if the project violates the second condition related to the progress towards critical loads, it is not permitted; otherwise it is accepted.

Apart from the administrative burden, this procedure based on bilateral trades may also not lead to a cost-effective allocation. The sequential order of trades would be very important indeed[7]. Whether a transaction is allowed or not may depend on whether or not it is proposed before some other transaction. However, trading can be seen as inducing Pareto improvements, provided third parties are not significantly affected. On the whole the initial allocation will be improved, if not made wholly cost-effective.

2.3 The exchange rate

The environmental impact of SO_2 emissions depends on the location of source and receptor. Hence, 'one unit increase from one source cannot be offset by one unit decrease from another

7. Linear programming models of least cost solutions implicitly assume that trading takes place in a multilateral simultaneous manner. In the real world, trading involves individual transactions occurring sequentially. The models may assume patterns of trade that may be impossible in reality. For example, a bilateral trade between two sources may result in target deposition being exceeded at a particular location, but this in turn may be offset by an additional trade with another source. Such a trading pattern would be allowed in the model but disallowed in reality as pollution targets are (temporarily) exceeded. The models may therefore overestimate the potential gains from trade. For a fuller discussion of this issue see Atkinson & Tietenberg (1991).

source. The exchange rate, also termed the offset rate, may be greater or smaller than one' (Førsund and Nævdal, 1994). The design of a trading scheme must respect this if it is to be compatible with deposition constraints. Two options are possible:

- to accept trading between trading zones on the basis of exchange rates fixed for the whole period.
- to refuse trading between trading zones, but to define trading zones as covering a large geographical area so as to offer wider trading opportunities and limit transaction and administrative costs.

In the first case, exchange rates should reflect the relative intensity of the marginal damage generated by one unit of emissions. However, it is difficult to find a workable rule that is able to reflect this requirement. First, marginal damage functions are not known and some proxy has to be used. Second, while a practical rule should keep its value through time, the conditions for optimality require a revision of exchange rates after each trade. Since emissions are concentrated in a limited number of sources, individual trades will generally have a non-negligible impact on the distribution of acid deposition. But continual adjustment of exchange rates would make trading unpredictable for agents, significantly complicate investment decisions and be administratively impractical.

The second solution sticks to a one to one exchange rate within trading zones considered to be homogeneous. This simple and robust approach may be viewed as more accessible, and easy to implement, being less dependent on central modelling and revision of information. However, it is not totally satisfying since the supposed homogeneity of each zone is an artificial construct to some extent.

The choice between these depends on both political judgement about the level of environmental guarantees offered and administrative practicality. Most proposals have explored exchange rates[8], without giving too much attention to guarantees of environmental improvement. Here we explore the second option as it may provide a higher level of guarantee. The key issue then becomes the size of the trading zone and the trade off this implies between environmental effectiveness and economic efficiency. This is discussed further in Section 3.

2.4 Which periodicity?

The issue here is whether agents should be allowed to trade continuously or whether trades should take place periodically through an organised mechanism. This is relevant at two levels: government-to-government trading and plant-to-plant trading. The best solution will depend upon the level.

8. For an allowance trading scheme based on exchange rates, see Amann et al. (1994).

With a two-level system of allowance trading, governments will trade on the basis of their national caps; once national caps have been decentralised to plants, the latter will trade together at the EU level. The intergovernmental market controlling the level of national caps should be made highly predictable for plants, in order to ensure the security of the allowances they receive from public authorities, and to allow them to engage in rational investment strategies. A clear means of providing this predictability would be to organise a discrete, periodic intergovernmental market (every 4 years, for instance), with advance transactions, i.e. transactions having effect some years later (say 3 years). This would mean that basic plants can move in a predictable institutional environment, with a secure horizon in the range of 3 to 7 years. For plant level trading the period could be much shorter - every six months or a year for example.

An alternative approach would be for governments to regulate the total quotas given to plants on a continuous base, as active operators on national markets. However, this approach could introduce instability and unpredictability into the market, or raise the fear that governments will behave in an arbitrary manner. These factors may turn out to be an obstacle to technological innovation, when the weight and sunk costs of industrial investments in desulphurisation equipment are taken into account.

3. ABOUT ZONING AND SCALING

The issue of zoning is a key one for the practicability of the trading schemes considered in this paper. The first problem to be addressed concerns the scale of trading zones, while the second concerns the criteria by which zones are defined.

At one extreme, we have grid-cells, i.e. relatively small territorial units of 150km by 150km. At the other extreme, we face one unique zone, the European territory covered by the LRTAP Convention. If trading is confined to specific zones, then larger scales provide more opportunities for improving economic efficiency, but this is achieved at the expense of lower levels of environmental security. Smaller scales offer greater environmental security but with fewer opportunities for profitable trade. Maintaining practical viability with sufficient potential for economic efficiency gains should be the relevant criteria for selecting the ‘best’ scale, not just having a complete guarantee about the environmental protection of every small part of the European territory. In this latter case, too much would be paid for environmental certainty. But how can we proceed in this direction? If the existing grid of 150km by 150km is to be used for trading, then allowing trade between all zones using a matrix of exchange rates is inescapable if sufficient flexibility is to be achieved. But this solution would not avoid hot spots and, if the exchange rates remain fixed, it would not provide the expected environmental guarantees. From this, it may seem preferable to stick to a one to one rate of exchange within homogeneous zones. This alternative requires the definition of a limited number of *macro zones* to give sources a sufficient margin of flexibility.

There are two main possibilities for defining such macro zones. The first consists in establishing a number of categories of exceedance of critical loads and to map the European territory according to these categories. In this case, two territorial units belonging to the same categories may not be adjacent. Allowing a one to one exchange rate within each equivalence class is appealing, since emissions will have a broadly similar effect on the environment. However, the risk of having hot-spots with an unduly large concentration of pollutants in some places cannot be excluded. Therefore it may be useful and prudent to introduce some additional restriction.

This may be provided by an alternative way of designing zoning - identifying homogeneous geographical zones; that is, zones having a geographical unity in terms of contiguity and at the same time the same broad level of excess deposition over critical loads. Several adjacent cells with similar sensitivity to deposition could be joined in a single macro zone. Such a grouping extends the trading possibilities between deposition allowances. Trades would only be allowed between sources having emissions falling in the same macro zone.

Such a grouping can be made revisable. Since the ultimate target is formulated in terms of respecting critical loads for each basic unit zone, progress in that direction may be supported by a transitional approach to the scaling of trading zones. The initial step would be organised on the basis of a limited number of macro zones. Such macro zones would not constrain sources enough to ensure everywhere compliance with critical loads targets. At later stages, these zones could be scaled-down.

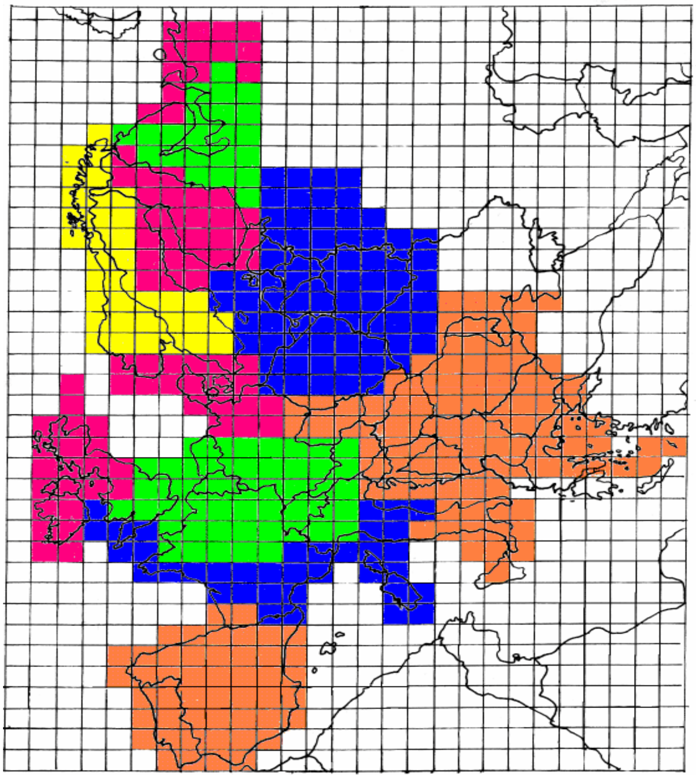
One critical point for the dynamics of scaling from an incentive viewpoint is that the future evolution of the grid should be communicated to participants well in advance so that they can develop strategic compliance plans incorporating early adaptation. Otherwise the outcome could be very inefficient. For instance, mistakes in capital investment might be induced. The authorities might therefore announce that the existing zones would be narrowed five years later, and then again ten years later. Such announcements of changes in the scale of trading zones will limit the problem of hot spots from the start, because for every decision having a medium or long term time horizon, specifically for planning investments (desulphurisation equipment, etc.), plant operators will have to take into account the announced changes. By the end of the process the long run targets fixed by the Oslo Protocol, i.e. respecting critical loads at the level of the grid-cell, will have to be met and this will give much less scope for trading.

With respect to practical matters, what type of zoning may reasonably be considered for an initial step? It seems that defining five trading macro zones in which critical loads are exceeded may make sense on both economic and ecological grounds. A recommendation proposed by Bailey, Gough and Millock (1994) considers such a grouping of unit zones having adjacent sensitivity. Five classes of acid sensitivity are used by them to classify each grid cell and achieve groupings accordingly.

Such a 'sensitivity' classification is not completely satisfying. Two areas being classified in the same sensitivity class may suffer unequal damage due to different levels of deposition: marginal damage not only depends on sensitivity levels but also on basic deposition received by zones in excess of critical loads. This is the reason why I suggest consideration of another type of zoning based on excess deposition over critical loads. With five classes of excess deposition, five critical macro-zones can be distinguished; they are surrounded by large areas where critical loads are not exceeded (see the Figures 2 and 3).

This mapping convincingly shows that drawing macro zones is not an entirely arbitrary exercise. By accepting some kind of 'sacrifice'[9] for a few cells, it is possible to identify homogeneous zones of a large scale. Meanwhile, significant areas of the European territory are relatively unaffected by acid deposition. The latter have depositions that do not exceed critical loads. The 60 per cent abatement constraint will not be binding for them. Countries like Spain, Greece and Portugal are broadly in this category.

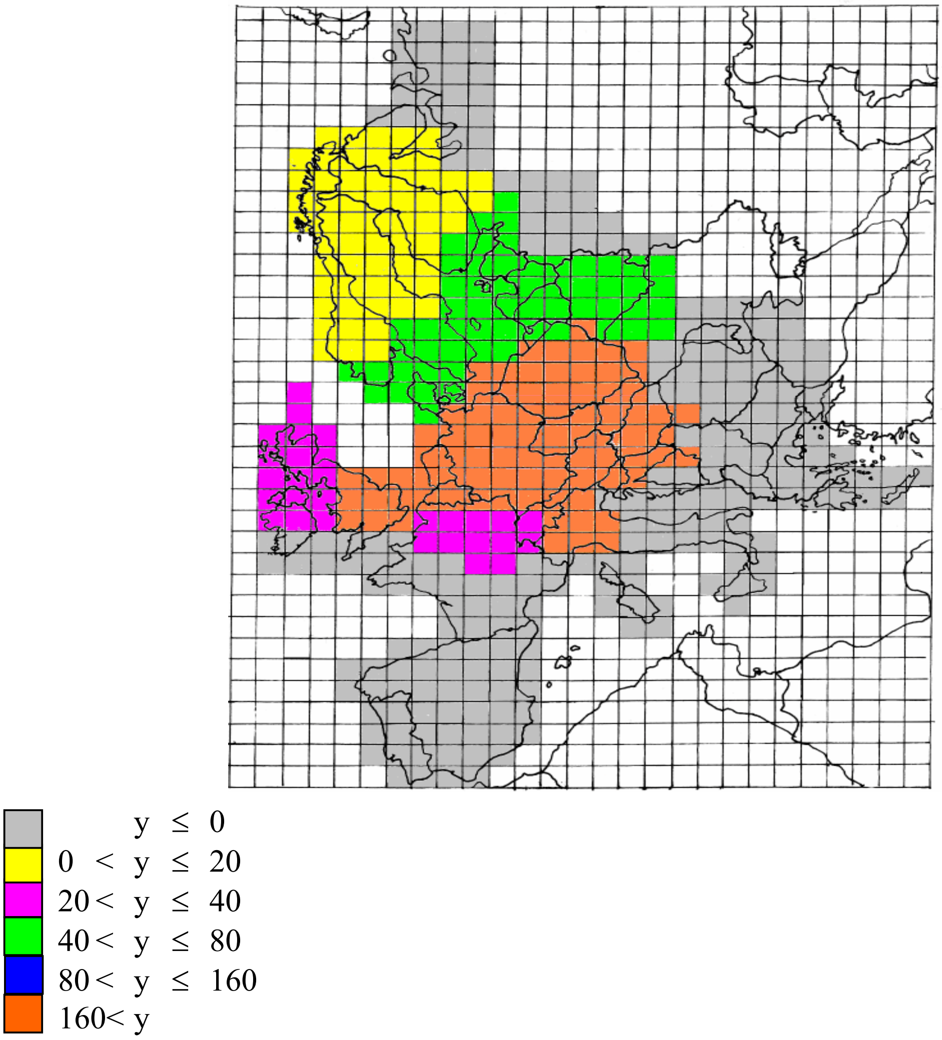
Figure 2: Aggregate zoning on the basis of equivalent ecological sensitivity



- (yellow) $x \leq 20$
- (magenta) $20 < x \leq 40$
- (green) $40 < x \leq 80$
- (blue) $80 < x \leq 160$
- (orange) $160 < x$

9. Macro-zoning does not exclude the risk that, in some areas, deposition may increase or may not decrease, though a significant decrease will be achieved in another part of the same macro-zone. Anyway, the target of a 60 per cent abatement rate of excess over critical loads does not directly refer to damage. In places where critical loads may be slightly exceeded, the target is the same as in places where the excess is of a greater magnitude.

Figure 3: Aggregate zoning on the basis of excess acid deposits over critical loads (1990 data)



4. IMPLEMENTING A SYSTEM OF TWO SIMULTANEOUS, COUPLED ALLOWANCE TRADING MECHANISMS

The trading system has to take into account two basic heterogeneous constraints: national emission caps and deposition limits in unit zones. Given this, a first possibility is to conceive of two different systems of trading, one for each constraint, which are coupled together to allow sources to emit a given amount of SO_2. To obtain the right to emit one tonne of SO_2, a source having been involved in emissions trading should possess one allowance of each type[10].

10. Plants not involved in emissions trading will just be submitted to the requirements of the emission allowances they receive from their governments. Additional deposition constraints become actual only when plants begin to trade. This reflects the view that trading will be politically accepted only if it provides both an economic improvement (cost abatement) and an environmental improvement.

With such a system, each source would then operate in two types of allowance market - one for *emissions allowances*, working at the EU level and based on initial allocations distributed by national governments, and one for *deposition allowances,* focused on the specific constraints for each deposition zone in Europe (including non EU regions) and based on initial allocations distributed by the central authorities (the UNECE Secretariat for example). This system would take account of emissions from sources which belong to the territory of the EU, but generate acid deposition outside the EU. In the framework of the Oslo Protocol, the same basic constraints operate whatever the territory involved.

For the market in deposition allowances, I have chosen to explore only one option, the one in which transactions are only authorised between sources generating deposition within the same deposition zones[11]. In that case, there will be as many markets of the second type as there are deposition zones. If there exist *n* unit zones in Europe, a source would have to operate on (at most) *n + 1* markets (the *n* deposition zone markets and the EU wide market of nationally allocated emissions allowances).

In order to couple the two types of allowances, a pollutant transportation model (such as EMEP[12]) has to be used for translating emissions into deposition or deposition into emissions. The basic framework for a system of this type is summarised in Table 1 and Figure 4.

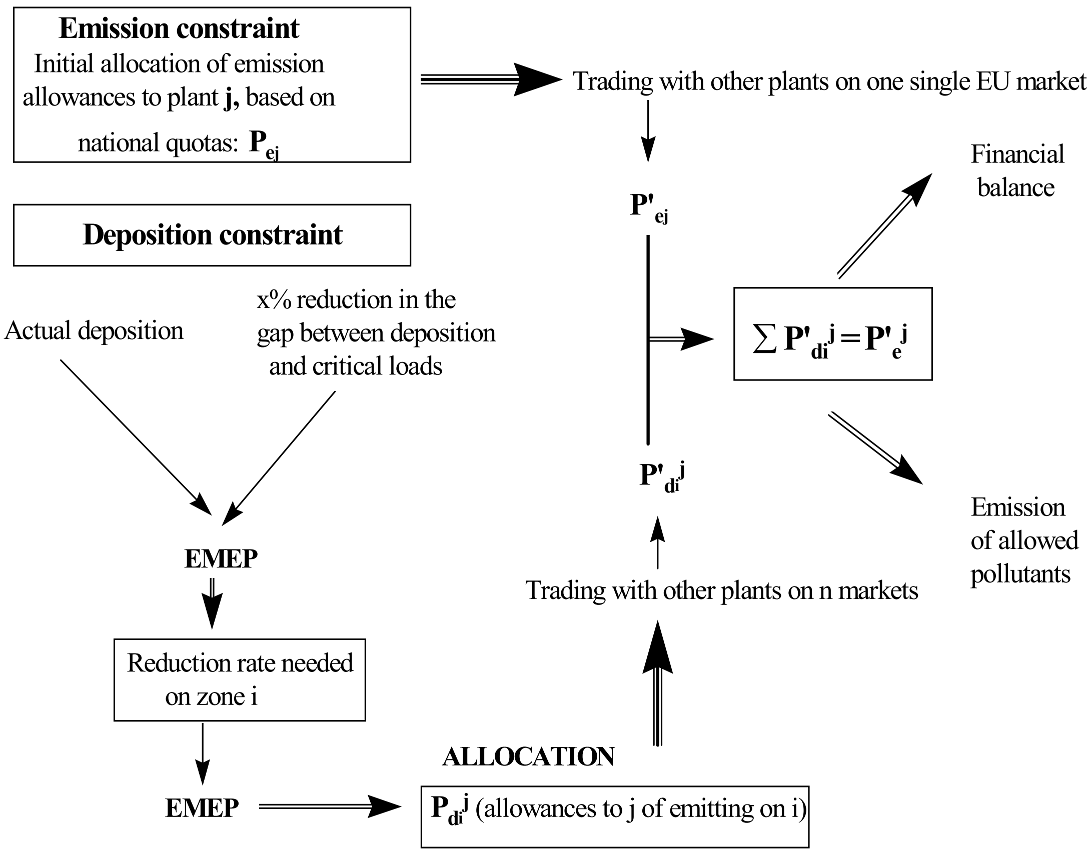
Table 1: A double system of allowances

	Emissions allowances	**Deposition allowances**
Units	tonnes	tonnes of deposition in zone *i*
Number of markets	1	Up to n - the number of unit zones
Trading	Unconstrained	only between sources generating deposition in the same zones
Allocation	national governments (political)	central authorities (using EMEP)

11. For example, if plant A has deposition in zones a, b, c, and plant B has deposition in zones b, c, d, then A and B can trade their respective depositions in zones b and c. But for this trading to be profitable, A will also have to trade with plant C or D having deposition in zone a, so as to meet all the constraints related to the zones on which emissions are deposited.

12. EMEP stands for the Co-operative Program for Monitoring and Evaluation of the Long Range Transmission of Air Pollutants in Europe. This is a subsidiary body to the CLRTAP and provides the official estimates of pollutant transportation and deposition within Europe.

Figure 4: A double system of allowances



This double system of allowance trading may be more practical that it seems at first sight. In the deposition markets, potential trading partners are more limited in number. They may be well-known to each potential participant. Thus, the problem of finding potential partners would be rather easy to overcome. At the same time, the market could be too thin, making it difficult to find partners ready for an exchange. The possibility of strategic interference amongst competitors (market power) cannot be avoided either. The importance of these difficulties may be expected to be proportional to the number of zones. Similarly the administrative practicality of the scheme is inversely proportional to the number of zones. These considerations strengthen the argument in favour of a small number of zones of large geographical area.

When potential traders know their trading opportunities on the deposition markets, they can adjust their strategy on the market for emissions. The emissions allowance market provides flexibility regarding the way the initial allocation has been dealt with politically by governments, although grandfathering is the most probable choice for political reasons, and provides large opportunities for exchange for suppliers and buyers. No specific additional constraints are necessary on this market since the deposition constraints are tackled by the deposition market.

This heterogeneous combination of two types of allowance can be seen as an incentive to trade. Participants need sufficient allowances of both types (emission and deposition) to be able to continue operating. If they have insufficient allowances of one type they must either buy more or reduce emissions. If they have more allowances of one type than they can use they have an incentive to sell. In each case, it is more rewarding to engage in trade than to stay in their present position.

While this scheme is significantly more complicated than, for example, the US Acid Rain Program, it may still be viable. The complications arise not from the scheme itself but from the dual constraint embodied in the existing UNECE regime[13]. Since these constraints are unlikely to be abandoned, it is necessary to devise a scheme that fully incorporates them. To minimise transaction costs and expand trading opportunities it will be necessary to define trading zones that are larger than the EMEP unit zones - even if this solution does not provide an absolute guarantee of environmental improvement for each unit zone.

The following sections describe two variants of an integrated allowance trading scheme. As in the previous scheme, physical modelling of pollutant transportation and deposition plays a central role, in both the allocation of allowances and the authorisation of trades. Proposed trades are tested with physical models for ascertaining their impacts on deposition for each unit zone. The variants differ as regards the rules for the initial allocation of allowances and the type of incentive mechanism incorporated to make reaching critical loads targets more attractive to plants (variant 1) and governments (variant 2).

5. AN INTEGRATED TRADING SYSTEM, WITH AN AUCTIONED MARKET FOR 'UNUSABLE ALLOWANCES'

Here, the initial allocation is organised in two steps. First, a potential allocation of 'emission' allowances to individual sources (plants) by national governments is calculated on the basis of national ceilings and allocation criteria chosen by those governments. The subsequent deposition from each source for each deposition zone is then assessed with the help of the EMEP model. In the meantime, the deposition target is used as the basis for a calculation of an overall deposition cap for each deposition zone. These zone deposition caps are then distributed proportionately to the sources responsible for the deposition, also using EMEP. This gives the potential 'deposition' allowances. The two allocations are translated in comparable terms (units of emissions) for each source by using the vector that describes how emissions from a source translate into deposition in the different deposition zones (say 10 per cent on R_1, 30 per cent on R_3, 25 per cent on R_{11}, and so on). Each source will have a different dispersion vector according to its location. At this moment, for each source, two different amounts of potential allowances are considered, the 'emission' one and the one derived from 'deposition'. The lower value of acceptable emissions is then selected, in order

13 It may be noted here that there are also dual constraints in the US system, but from different origins. Power stations must comply with both the federal Acid Rain Program and state regulations on local air quality. In this case, local regulations, for instance in the Mid-West, frequently commanded investments in scrubbers, leaving important amount of unused allowances available for sale. Symmetrically, after the achievement of such investments, those utilities were uninterested in purchasing allowances.

to satisfy the more binding constraint. On this basis an *actual* quota of allowances is allocated to the sources. These may be termed *usable and tradable allowances*. They can either be used directly, to cover actual emissions of the source, or traded, if sources take measures to abate their emissions under this quota. Each allowance of one ton of SO_2 is then defined as a vector of deposition in *n* zones.

At the same time, individual sources (plants) are given an extra amount of potential allowances responding to the difference between the less binding constraint (the emission or deposition, it depends on the source location) and the more binding one. This extra amount cannot be used directly but may be used for trading. These can be termed *unusable, tradable allowances*. This extra allocation will supplement the basic entitlement and give an additional incentive to sources to enter into allowance trading, since they can benefit from trading opportunities which remain compatible with both constraints of the regime. The operation of the mechanism could be as follows:

- Just like 'usable and tradable allowances', unusable ones are defined as a deposition vector for one ton of SO_2. Two cases can be considered.

  * Whenever 'unusable and tradable allowances' derive from a more binding emission constraint they can be sold freely but they can only be used by the buyer for a use touching the same deposition zones in Europe as the ones that would have been affected if the seller had used them directly. This means that such transactions have to be checked as regards the deposition zones affected. For instance, if 100 tonnes of unusable allowances are sold by a source to another, what is really sold is a deposition right reflecting the structure of deposition of the seller, say 20 tonnes in Z_1, 30 tonnes in Z_2, 50 tonnes in Z_3. So the entitlement obtained through the purchase of 100 'unusable deposition allowances' is a vector $D_{1,2,3}$ (20, 30, 50). It is possible that the buyer cannot use the whole spectrum of what it buys, due to its own different location and different structure of deposition from its emissions.

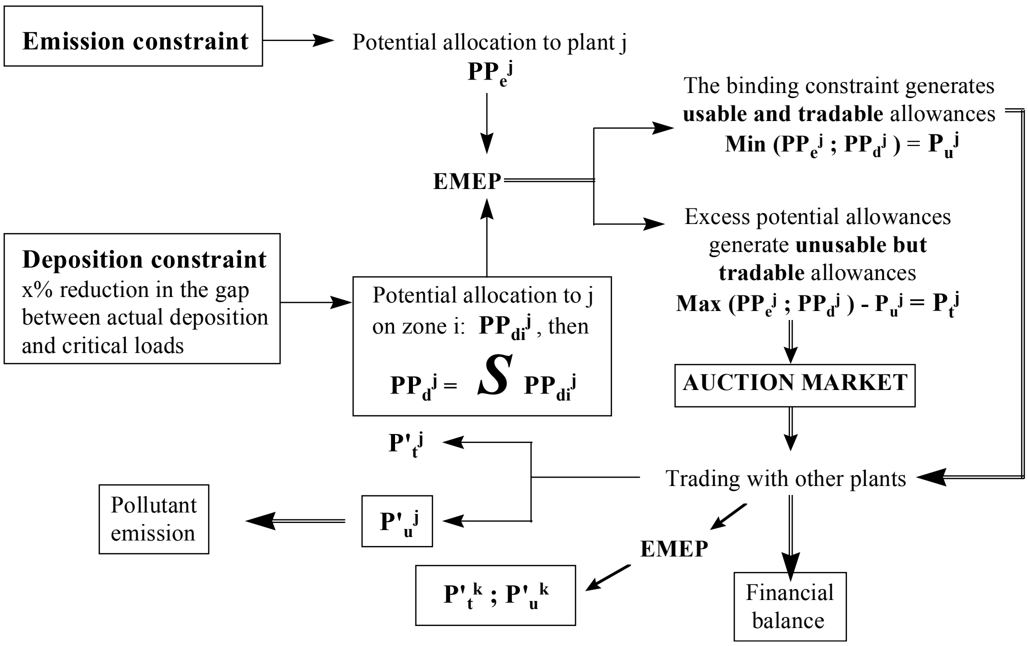
  * Whenever unusable allowances are related to a non binding emission constraint, the deposition vector for one ton of SO_2 will be zero, limiting the possible usage of such allowances to the cases when users need to complement quotas limited by an emission constraint and not a deposition one.

- So, the amount of 'unusable, tradable allowances' may evolve with time, following the various transactions. At any moment, the net amount of 'usable allowances' is defined by the level of the most binding constraint (emission or deposition), and the amount of 'unusable allowances' can be calculated as the difference between the two potential allowances (emission and deposition).

- A financial mechanism could be set up to facilitate the valuation of 'unusable allowances' on a market. They may feed an auction market organised by the authorities on behalf of sources. The revenues raised by the auctions could be refunded to the source entitled to it[14]. This 'last resort market' would be open to any source, but the buyers will be subject to the same constraints of usage that were previously described for all sources.

This system may seem complicated, but once an algorithm has been defined for making calculations plants should be able to deal with the system. The structure of the system is summarised in figure 5:

Figure 5: An integrated system with an auctioned secondary market



With such a system, an auction market for allowances is generated in addition to bilateral trades. This has the following advantages:

- It gives additional flexibility and safety to sources and avoids the strategic retention of allowances. The mechanism would be similar to auctions organised by EPA in

14. This is termed a zero revenue auction

the US; any source looking for allowances and not finding them through bilateral trade could enter this recourse market.

- It facilitates the emergence of a public reference price for SO_2 allowances, and allows comparisons to be made between the various national markets. This compares with bilateral trades which are normally private, with no release of price information.

In providing a means to make economic use of ‘unusable allowances’ this integrated system creates an incentive for most countries and individual sources to accept the constraint of the percentage critical loads target, since going further than national ceilings would be compensated for in this way. This may enhance the political acceptability of the regime and accelerate progress towards the target of respecting critical loads. In contrast, if no opportunity to benefit from unusable allowances was given, the political and logical coherence of this integrated solution would be open to question. This is because it would give a strong weight to the critical loads goal without providing the incentives necessary to achieve this goal.

6. AN INTEGRATED SYSTEM INCORPORATING COMPENSATIONS FOR STATES

This procedure for the initial allocation reverses the previous solution. One begins by considering current emissions from sources and simulates, using EMEP, the subsequent deposition in each basic deposition zone. Two cases are then possible. If the critical loads are not exceeded, the source is credited with a ‘deposition allowance’ corresponding to its current emissions. If the critical loads are exceeded, the current target of a percentage closure of the gap between deposition and critical loads is used to calculate a deposition cap for the deposition zone; this cap is allocated proportionately to each source having deposition in the zone. This defines the *first* formula for determining the potential SO_2 deposition allowances to be received by each individual source. Consider this example: one source S_1 has three tonnes of deposition on a deposition zone Z_1, for which the cap is not exceeded, and eight tonnes of deposition on Z_2 where the cap is exceeded. Then S_1 will first receive 3 P_{d_1}. If Z_2 is receiving a total amount of excess deposition of 20 tonnes and responsibility of S_1 for this is 5 per cent, then (assuming a deposition target of 60%) it will also receive:

$$\{8 - [(20 \text{ x } 60\ \%) \text{ x } 5\ \%]\} = 7.4\ P_{d_2}.$$

So the first potential deposition allocation of S_1 is: $P^1{}_D = 3\ P^1{}_{d_1} + 7.4\ P^1{}_{d_2}$

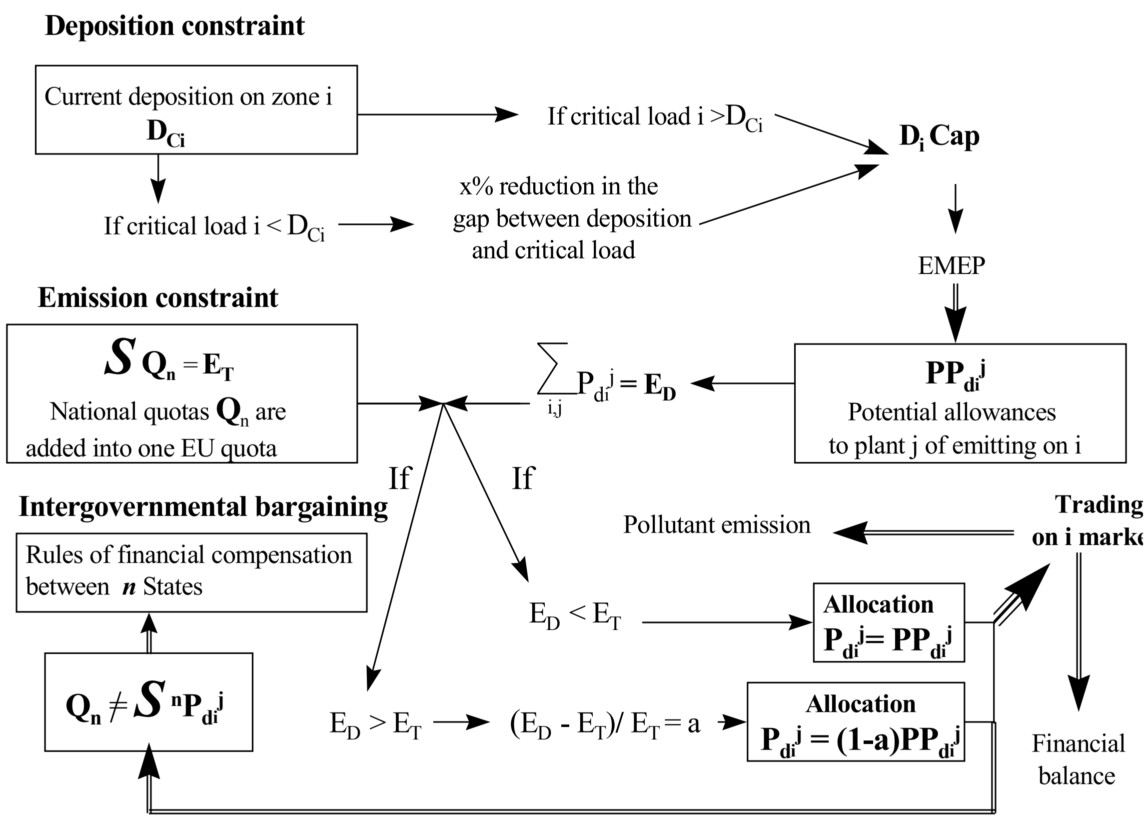
Then the total amount, E_D, of such SO_2 allowances given to EU sources is calculated to test the compatibility of this allocation with the Oslo Protocol, as regards abatement commitments expressed in national ceilings:

$$E_D = \sum_{i,j} P^{i}_{d_j} \text{ for each source } i \text{ and zone } j$$

An EU cap on emissions, E_T, is also calculated as the sum of SO_2 emissions compatible with agreed national ceilings. If $E_D \leq E_T$, the first allocation is actually implemented for sources, since it satisfies at the same time the total EU cap and the critical loads target for each European zone. If on the contrary $E_D > E_T$, some additional restriction is needed. It can be argued that a proportional reduction on the *first* formula of allocation of allowances will be the appropriate solution for every individual source. For instance, if $E_D = 1.2 \times E_T$, the actual initial allocation of S_1 will be:

$$P^1_T = [(3/1.2\, P_{d_1} ; 7.4/1.2\, P_{d_2})] = (2.5\, P_{d_1} ; 6.15\, P_{d_2})$$

Figure 6: An integrated system with compensations for States



Such a procedure fits an integrated EU political context, since member countries are required to transfer their national quotas to the EU, so as to obtain a global EU wide cap. The prominent role given to the critical loads targets also fits this context. But under what conditions will this solution be acceptable to governments? Some countries will certainly see their actual quota reduced when compared to the agreed national ceilings in the Oslo Protocol or to the first variant considered above. It seems quite natural to envisage some mechanism of financial compensation. Governments, not sources, are proposed to be compensated, since governments have to be convinced to accept additional restrictions.

What basis can be imagined for this compensation? Since the breakthrough of the Oslo Protocol was only possible because the parties accepted some integrated assessment as a basis for elaborating an optimal international plan, parties could agree to calculate compensation using the same tool. The compensations table could then be calculated as the difference in national costs resulting from two allocations: the agreed national ceilings of the Protocol and the allocation resulting from the procedure that has been just described.

7. CONCLUSION

The three proposals examined here were specifically conceived to address the two major constraints by which the institutional context of the EU and UNECE SO_2 requirements have been interpreted: national emission ceilings and deposition constraints for geographical zones. What can be the future of such proposals, if they are to become reality? Can we imagine an evolution towards a system of tradable allowances structured by a single constraint, i.e. deposition allowances?

To achieve an evolution towards a deposition allowance trading scheme, two conditions need to be met:

- all countries of the EU must abandon their national quotas to a common EU sovereign body, for redistribution according to some agreed rule (the political condition);
- constraints related to deposition allowances should be more binding, in every area, than the ones related to the emissions allowances (the technical condition).

If, for the sake of discussion, we take the first political condition as met, achievement of the second one looks rather doubtful. Deposition markets would frequently be the more binding ones, but this will not always be the case, since all EU countries have caps on their emissions, but critical loads are not exceeded in significant part of the EU territory. A double system takes into account different rationales for allocation, which is an attractive property for achieving socio-political acceptability for the trading regime. Moreover the two types of allowances do not cover the same territory: emissions allowances, according to the proposed schemes, would be limited to the EU territory, while deposition allowances will be of concern for all Europe, including non EU deposition zones which are affected by EU emissions.

Evolution towards a one-constrained scheme would require a distribution of power in the EU which is rather different from that at present, where nation states have preserved significant political autonomy. The political equilibrium reached with the Oslo Protocol may be redefined in the future. In the meantime, the type of mechanism that has been presented in this paper does fit well in this framework. Either a double market system or an integrated one provides the sort of equilibrium needed and may operate for a long time as an intermediate tool between national approaches, as reflected in national ceilings, and the pure co-operative European approach reflected in the critical loads reference.

REFERENCES

Amann, M., F. Førsund and G. Klaassen (1994), 'Emission Trading in Europe with an Exchange Rate', *Environmental and Resource Economics*, **4**, pp. 305-330

Atkinson, S. and T.Tietenberg (1991) 'Market failure in incentive-based regulation: the case of emissions trading', *Journal of Environmental Economics and Management,* **21**, pp. 17-31

Bailey, P., C.Gough and K.Millock (1994), 'Sulphur emission exchange rates', in F.Førsund and G. Klaassen (eds), *Economic Instruments for Air Pollution Control*, Kluwer Academic Publishers.

Cros, C. and O. Godard (1998), 'The economic design of a potential tradable permit system for SO2 emissions in the European Union', *European Economy,* 'Getting Environmental Policy Right. The Rational Design of European Environmental Policy from an Economic Perspective', Report and Studies (Vol. 1), forthcoming Spring.

Godard, O. (1995), 'Trajectoires institutionnelles et choix d'instruments pour les politiques d'environnement dans les économies en transition'*, Revue d'études comparatives Est-Ouest*, **26**, (2), CNRS, juin, pp. 39-58.

van Ierland, E.C., Kruitwagen,S. and Hendrix, E.M.T. (1994), 'Tradable Discharge Permits for Acidifying Emissions in Europe: 'guided' bilateral trade', in F.Førsund and G. Klaassen (eds), *Economic Instruments for Air Pollution Control*, Kluwer Academic Publishers.

Klaassen, G. (1996), *Acid Rain and Environmental Degradation. The Economics of Emission Trading*, Aldershot, UK and Brookfield, US: Edward Elgar.